We have seen that under some conditions, the binomial PMF is well approximated by a Poisson PMF.

But we have also seen the central limit theorem that tells us that the binomial PMF can be approximated using a normal random variable.

Can we reconcile these two facts?

Let's look into the situation in some more detail.

Consider a Poisson process that has rate equal to 1.

And consider that Poisson process running over the unit time interval.

We take the unit interval, and we split it into many small sub-intervals, where each sub-interval has a length of 1/n.

And let Xi be the number of arrivals that we get during the i'th interval.

Xi is a Poisson random variable.

And the mean of that random variable, or the parameter of that random variable, is just the duration of the time interval, since the rate is 1.

So it's a Poisson random variable, with parameter 1/n.

Now, let us look at the total number of arrivals.

The total number of arrivals is the sum of how many arrivals we had during each one of these intervals.

And we know the distribution of S. S is a Poisson random variable, with parameter equal to 1.

Now, here what we have is a sum of random variables that are independent and identically distributed.

They are identically distributed, because all of these intervals have the same length.

And they're independent, because in the Poisson process, what happens in different intervals are independent events.

So we are in a situation where we could apply the central limit theorem.

We have a sum of many independent, identically distributed random variables.

And by letting n go to infinity, the central limit theorem appears to tell us that S is going to be normal.

Now, how can we reconcile these two facts?

We know that the Poisson distribution is not the same as a normal distribution.

What is the catch?

Well, the catch is the following-- the central limit theorem applies to a situation where we fix a certain probability distribution, the distribution of the Xi's.

And it tells us that as we add more and more of these Xi's, asymptotically, we obtain a distribution that's well approximated by a normal.

On the other hand, what we have here is actually different.

The Xi's do not have a fixed distribution.

But rather, the distribution of Xi depends on n.

That is, if we change n so that we're adding more random variables, we're adding more random variables that are now coming from a different distribution.

And this is not a situation to which the central limit theorem applies.

And therefore, this conclusion here is not justified.

And so there's no contradiction between the two types of approximations.

To summarize, the situation is as follows.

Consider a binomial random variable with some parameters n and p.

Now, let p be fixed.

And let n go to infinity.

In that case, the binomial random variable can be thought of as the sum of n Bernoulli random variables.

And those Bernoulli random variables have a parameter p, which is fixed.

So we're dealing with the sum of iid random variables from a fixed distribution.

And this is the situation where the central limit theorem applies.

And we have a normal approximation.

On the other hand, if we take the product n times p, which is the expected value of this binomial, to stay constant, but we let n go to infinity and at the same time let p go to 0, then in this regime, in the limit, this random variable will be well approximated by a Poisson random variable.

So we have two different approximations.

Both of them are valid, but they're valid in different regimes.

Now, although they're different, there's actually an interesting case in which the two will not really differ.

And this is the following.

Consider a Poisson random variable with parameter n.

And we're interested in the limit as n goes to infinity.

We can think of this random variable as the number of arrivals during an interval of length n in a Poisson process with arrival rate equal to 1.

Now, let's take this interval and split it into n intervals, each of which has a length of 1.

And let us call Xi the number of arrivals in the i'th interval.

Our Poisson random variable is going to be, of course, equal to the sum of the number of arrivals during each one of the intervals.

Each one of these random variables is Poisson with parameter equal to 1.

And these random variables are actually iid.

Now, what's happening in this case is that even when we increase n, because we're using constant length intervals, the distribution of the Xi's doesn't change.

So this is a situation in which we're going to get approximately a normal random variable as n goes to infinity.

So what we see is that a Poisson random variable, but with a very large parameter, starts to approach the normal distribution.

And this in particular will tell us that these two approximations that we have, in some regime, they would start to

agree.

Now, all of this discussion here has been asymptotic.

We talk about p going to 0 or n going to infinity.

But in any real situation, you will be given actual numbers.

And you cannot really tell, is this number close to 0, or is it not?

Here, we need some rules of thumb or maybe some experience.

Let's look at some examples.

In this case, n times p is equal to 1.

So the number of arrivals or the values of the binomial random variable will take values 0, 1, 2, 3, but with high probability, not a lot more than that.

So the binomial random variable is really a discrete random variable.

There's no way to approximate it with a normal.

On the other hand, p is very small.

So a Poisson approximation would be very reasonable in this situation.

On the other hand, if p is equal to 1/3, then definitely 1/3 is not a small number.

A Poisson approximation would not apply.

But n is pretty big.

So that a normal approximation would be appropriate.

And finally, in this case, we would obtain a Poisson approximation with parameter 100, because n times p is 100.

But a Poisson random variable with parameter 100 is also well approximated by a normal.

Or to think about it differently, we start with a Bernoulli distribution that's very skewed, [the] probability of success is just 100.

And this makes it difficult for the central limit theorem to apply when you start with a very asymmetric distribution.

On the other hand, because we're adding so many of them, the central limit theorem actually does take hold.

And so this is an example where both approximations will be valid.

So finally, to conclude, we have two different approximations.

They're valid in different regimes.

And in practice, you need to do some thinking before deciding to choose one versus the other approximation.